REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE: 31 December 1995

3. REPORT TYPE AND DATES COVERED: Final 01 January 1993-31 December 1995

4. TITLE AND SUBTITLE: Characterization of Transients

5. FUNDING NUMBERS: G-F49620-93-1-0102

6. AUTHOR(S): Stephane Mallat

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES):
New York University
251 Mercer Street
New York, NY 10012

AFOSR-TR-96

0374

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES):
Air Force
110 Duncan Avenue Room B115
Bolling AFB DC 20332-8080
ATTN: Dr. Jon Sjogren

nm

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

19960726 057

12a. DISTRIBUTION/AVAILABILITY STATEMENT: Approved for public release;distribution unlimited

13. ABSTRACT (Maximum 200 words) We introduced a signal representations to detect and analyze transients. A new algorithm called matching pursuit was developed to decompense any signal as a sum of waveforms that are chosen adaptively from a redundant dictionary of patterns, in order to best match the signal structures. This algorithm was applied to dictionary of dilated Gabor functions in order to characterize oscillatory transients of various sizes and frequencies. The asymptotic properties of this algorithm have been analyzed and we proved the existence of an attractor. This led to a general noise removal procedure which has been applied to audio signals. A fast matching pursuit algorithm was also designed and implemented in a software that is freely available on the internet. The matching pursuit algorithm has been extended in two dimensions for image processing. The image dictionary is composed of translated, dilated, and rotated wavelets. Applications to texture discrimination have been studied. To isolate patterns whose support may intersect, we have introduced a high resolution pursuit algorithm which was used to decompose high resolution radar signals.

14. SUBJECT TERMS: Wavelets, Target Detection, Adaptive Representation

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

1 Introduction

The complexity of structures encountered in ATR signals require to develop adaptive low-level representations. Although these signals are entirely characterized by their decomposition in a basis, a basis is a minimal set of vectors that is not rich enough to represent efficiently all components. Some signal structures are diffused across many basis elements and are thus difficult to analyze from this expansion. For example, image variations corresponding to edges and textures are not efficiently represented by the same types of waveforms. The same issue appears in sounds that includes transients that are well represented by short waveforms, and harmonics that are more efficiently decomposed over long waveforms with short frequency support. Instead of decomposing all signals over the same family of waveforms, we introduced an adaptive transforms choose the decomposition vectors depending upon the signal properties. These vectors are selected among a family of waveforms that is much larger than a basis, which is called a dictionary.

In this study we have developed an algorithm called matching pursuit algorithm (section 2), which decomposes signals over dictionary vectors that are selected with a greedy strategy. Most of the signal energy can be approximated with few dictionary vectors, which can be interpreted as essential signal features. The application of matching pursuit to sounds (section 3) and images (section 5) have been developed with dictionaries of time-frequency atoms and wavelets. The asymptotic properties of the pursuit have shown the existence of a chaotic attractor that we used for noise removal (section chaos). To isolate features whose support overlap, in collaboration with Alan Willsky group at MIT, we have developed a high resolution pursuit algorithm which can segment features closely spaced (section 7).

The extraction of information from signals also requires to analyze structures that are better modeled in a stochastic framework. For stochastic signals, we are not interested by the exact behavior of a particular realization but we want to identify the underlined process. One or few realizations give very little information about the underlying process. We thus concentrate on second order moment properties. A new algorithm to estimate the covariance of non-stationary processes has been introduced in collaboration with Pfr. Papanicolaou.

By iterating this decomposition up to the order m, we can decompose f into the telescoping sum

$$f = \sum_{n=0}^{m-1} \left(R^n f - R^{n+1} f \right) + R^m f. \tag{7}$$

Equation (5) yields

$$f = \sum_{n=0}^{m-1} < R^n f, g_{\gamma_n} > g_{\gamma_n} + R^m f. \tag{8}$$

Similarly, we write $\|f\|^2$ as a telescoping sum

$$\|f\|^2 = \sum_{n=0}^{m-1} \left(\|R^n f\|^2 - \|R^{n+1} f\|^2 \right) + \|R^m f\|^2 \tag{9}$$

which we combine with (6) to obtain an energy conservation equation

$$\|f\|^2 = \sum_{n=0}^{m-1} |< R^n f, g_{\gamma_n} >|^2 + \|R^m f\|^2. \tag{10}$$

A matching pursuit decompose any f into a sum of dictionary elements which are chosen to best match its residues. Although this decomposition is non-linear, we maintain an energy conservation as though it was a linear, orthogonal decomposition. An important issue was to understand the behavior of the residue $R^m f$ when m increases. We proved [1] that the matching pursuit converges, even in infinite dimensional spaces.

Theorem 1 *Let $f \in \mathbf{H}$. The residue $R^m f$ defined by the induction equation (5) satisfies*

$$\lim_{m \to +\infty} \|R^m f\| = 0. \tag{11}$$

Hence

$$f = \sum_{n=0}^{+\infty} < R^n f, g_{\gamma_n} > g_{\gamma_n}, \tag{12}$$

and

$$\|f\|^2 = \sum_{n=0}^{+\infty} |< R^n f, g_{\gamma_n} >|^2. \tag{13}$$

When $\mathbf{H}$ is of finite dimension, $||R^m f||$ decays exponentially to zero.

This theorem proves that any vector f is characterized by the double sequence $(< R^n f, g_{\gamma_n} >, \gamma_n)_{n\in\mathbf{N}}$, which specifies the expansion coefficients and the index of each chosen vector within the dictionary. After m iterations, (8) shows that the approximation error is

$$R^m f = f - \sum_{n=0}^{m-1} < R^n f, g_{\gamma_n} > g_{\gamma_n}. \tag{14}$$

The best approximation of f as a a linear expansion of $\{g_{\gamma_n}\}_{0\leq n<m}$ is the orthogonal projection of f on the space generated by this family of vectors. In general the vectors $\{g_{\gamma_n}\}_{0\leq n<m}$ are not orthogonal so the matching pursuit expansion is not equal to the orthogonal projection of f. An improved approximation was introduced in collaboration with Geoff Davis, by orthogonalizing the family $\{g_{\gamma_n}\}_{0\leq n<m}$ with a Gram-Schmidt procedure and computing the orthogonal projection of f [4]. Such an orthogonal pursuit gives the better approximations at the cost of an increase computational complexity.

3 Sound Pursuit

To analyze the time and frequency localization properties of one-dimensional oscillatory signals such as speech, Zhifeng Zhang [1] used a large dictionary of time-frequency atoms. Our signal space is $\mathbf{L}^2(\mathbf{R})$ and we construct such a dictionary by scaling, translating and modulating a single window function $g(t) \in \mathbf{L}^2(\mathbf{R})$. We suppose that $g(t)$ is an even and real function of unit norm. For any scale $s > 0$, frequency modulation ξ and translation u, we denote $\gamma = (s, u, \xi)$ and define

$$g_\gamma(t) = \frac{1}{\sqrt{s}} g(\frac{t-u}{s}) e^{i\xi t}. \tag{15}$$

The index γ is an element of the set $\Gamma = \mathbf{R}^+ \times \mathbf{R}^2$. The factor $\frac{1}{\sqrt{s}}$ normalizes to 1 the norm of $g_\gamma(t)$. The function $g_\gamma(t)$ is centered at the abscissa u and its energy is concentrated in a neighborhood of u, whose size is proportional to s. Let $\hat{g}(\omega)$ be the Fourier transform of $g(t)$. Equation (15) yields

$$\hat{g}_\gamma(\omega) = \sqrt{s}\hat{g}(s(\omega - \xi)) e^{-i(\omega-\xi)u}. \tag{16}$$

Since $|\hat{g}(\omega)|$ is even, $|\hat{g}_\gamma(\omega)|$ is centered at the frequency $\omega = \xi$. Its energy is concentrated in a neighborhood of ξ, whose size is proportional to $\frac{1}{s}$. The

dictionary of time-frequency atoms $\mathcal{D} = \{g_\gamma(t)\}_{\gamma\in\Gamma}$ is a very redundant set of functions in $\mathbf{L}^2(\mathbf{R})$ that includes window Fourier frames and wavelet frames.

A matching pursuit chooses the time-frequency atoms of $\mathcal{D}$ which are "best" adapted to expand f. Since a time-frequency atom dictionary is complete, Theorem 1 proves that

$$f = \sum_{n=0}^{+\infty} < R^n f, g_{\gamma_n} > g_{\gamma_n}, \tag{17}$$

where $\gamma_n = (s_n, u_n, \xi_n)$ and

$$g_{\gamma_n}(t) = \frac{1}{\sqrt{s}_n} g(\frac{t-u_n}{s_n}) e^{i\xi_n t}. \tag{18}$$

Numerical study have been performed on speech and audio signals [1], [4].

4 Chaotic Attractor and Noise Removal

The asymptotic convergence of a matching pursuit has further been studied by analyzing the behavior of the normalized residue

$$\tilde{R}^n f = \frac{R^n f}{||R^n f||}.$$

The non-linear map defined by $\tilde{R}^{n+1} f = \mathcal{M}(\tilde{R}^n f)$ exhibits chaotic properties. Experimental data suggest that the normalized residues of a normalized pursuit converge to a chaotic attractor. In low-dimensional spaces, Geoff Davis [3] proved that $\mathcal{M}$ is topologically equivalent to a left-shift map operator, whose chaotic properties are entirely known. In high dimensional spaces, the analysis was performed for a particular dictionary of Diracs and complex exponentials. Numerically, one can observe that the first few iterations of the pursuit extracts the components of f which are strongly correlated with dictionary vectors, which we call coherent part. The remaining residue does not correlate strongly to any dictionary vectors and its properties depend upon the attractor of the chaotic map. We call it a "dictionary noise" For dictionaries of time-frequency atoms these residues converge to realizations of white noises [3].

By tracking the convergence of the residue to the dictionary noise attractor, we can isolate the coherent signal components that are well approximated by few dictionary vectors. Applications to noise removal from speech signals have been studied by Geoff Davis [3].

5 Image Pursuit

For image processing, we must select a dictionary that can characterize the local scale and orientation of the image variations. For this purpose, Francois Bergeaud [7] introduced a dictionary composed of several two-dimensional wavelets that have specific orientation selectivity. These wavelets are all derived from a two-dimensional window $g(x, y)$ that is modulated at a fixed frequency ω_0 along several directions specified by an angle θ in the (x, y) plane

$$g_\theta(t) = g(x, y) \exp[i(x \cos\theta + y \sin\theta)].$$

These oriented wavelets are then scaled by s and translated to define a whole family of wavelets $\{g_\gamma\}_{\gamma\in\Gamma}$ with:

$$g_\gamma(x, y) = \frac{1}{s} g_\theta(\frac{x-u}{s}, \frac{y-v}{s}). \tag{19}$$

The multi-parameters index $\gamma = (\theta, s, u, v)$ carries the orientation, scale and position of the corresponding wavelet.

In numerical computations, the scale is restricted $\{2^j\}_{j\in\mathbf{Z}}$ and the angles are discretized. This wavelet dictionary used in the numerical examples include 8 orientations. The matching pursuit algorithm applied to this wavelet dictionary selects iteratively the wavelets, whose scales, orientations and positions best match the local image variations. Applications to texture discrimination and noise removal have been developed [7].

6 Fast Numerical Computations

At a first glance, a matching pursuit seems to require a hopeless amount of computations. These computations can however be considerably reduced with an efficient algorithm that prunes the dictionary with local maxima [1], [7]. For $f \in \mathbf{H}$, we call a local maxima in the parameter space $\mathbf{\Gamma}$ an index γ_0 such that for all γ in a neighborhood of γ_0 in $\mathbf{\Gamma}$

$$|<f, g_\gamma>| \leq |<f, g_{\gamma_0}>|. \tag{20}$$

For example, in a Gabor dictionary of one-dimensional time frequency atoms, the local maxima are computed for fixed scale. For each scale s, the local maxima are defined as indexes $\gamma_0 = (s, u_0, \xi_0)$ such that (20) is valid for any $\gamma = (s, u, \xi)$ with (u, ξ) in a neighborhood of (u_0, ξ_0).

At the step 1 of the algorithm we prune the dictionary with a local maxima selection. All inner products $\{< f, g_\gamma >\}_{\gamma \in \mathbf{\Gamma}}$ are computed. We choose a threshold ϵ and select only the local maxima that are large enough

$$| < f, g_\gamma > | \geq \epsilon ||f||.$$

The matching pursuit is then computed by induction as follow.

Suppose that the first n vectors $\{g_{\gamma_k}\}_{0 \leq k < n}$ have been selected. We denote by $\mathbf{\Gamma}_n$ the indexes γ such that $| < f, g_\gamma > |$ is a local maxima and $| < R^n f, g_{\gamma_0} > | \geq \epsilon ||f||$. We find g_{γ_n} which correlates $R^n f$ at best in this reduced dictionary

$$| < R^n f, g_{\gamma_n} > | = \sup_{\gamma \in \mathbf{\Gamma}_n} | < R^n f, g_\gamma > |.$$

We compute the inner product of the new residue $R^{n+1} f$ with all $\{g_\gamma\}_{\gamma \in \mathbf{\Gamma}_n}$ with an updating formula derived from equation (5)

$$< R^{n+1} f, g_\gamma > = < R^n f, g_\gamma > - < R^n f, g_{\gamma_n} > < g_{\gamma_n}, g_\gamma > . \tag{21}$$

Since we previously stored $< R^n f, g_\gamma >$ and $< R^n f, g_{\gamma_n} >$, this update is obtained in O(1) operations if the value $< g_{\gamma_n}, g_\gamma >$ can be retrieved in $O(1)$ operations. This is the case for the Gabor dictionary of one-dimensional time-frequency atoms and the dictionary of two-dimensional wavelets. The vectors in these dictionaries have a sparse interaction which means that for most $\gamma \in \mathbf{\Gamma}_n$, we have $< g_{\gamma_n}, g_\gamma > = 0$. There are thus few indexes γ for which the value of $< R^n f, g_\gamma >$ must be updated. The dictionary is further pruned by suppressing from $\mathbf{\Gamma}_n$ all indexes γ such that $| < R^{n+1} f, g_\gamma > | < \epsilon ||f||$. The iteration is then continued on this new index set $\mathbf{\Gamma}_{n+1}$.

If we iterate this procedure, the index $\mathbf{\Gamma}_n$ is progressively reduced until it gets empty for $n = m$. We then come back to the step 1 and replace f by $R^m f$. The local maxima of $< R^m f, g_\gamma >$ are computed and are thresholded with the new value $\epsilon ||R^m f||$. The pursuit is then continued on these maxima with the iteration previously described, until the index set is again empty for $n = p$. We come back again to step 1 by replacing f by $R^p f$ and continue the iterations. A software implementing matching pursuit for time-frequency dictionaries is available through anonymous ftp at the address cs.nyu.edu . Instructions are in the file README of the directory /pub/wave/software.

7 High Resolution Pursuit

The matching pursuit is a greedy strategy which does not use any "look-ahead" for selecting the dictionary vectors. When the features have a support that intersect, this can induce a selection of dictionary vector that is not optimal. In collaboration with Pfr. Alan Willsky group at MIT, we developed a high resolution pursuit algorithm that uses the same greedy strategy but which replaces the optimization of and $\mathbf{L}^2(\mathbf{R})$ correlation by a non-linear measure which is more sensitive to the local fit of features [9].

The high resolution pursuit algorithm was developed for two types of dictionaries. We used a dictionary of dilated and translated box spline functions to decompose high resolution radar signals. These features are then used for classification. We also developed an application to the detection of transitions in sounds by using a dictionary of wavepackets [9].

8 Estimation of Covariance

For general non-stationary processes the covariance matrix cannot be estimated reliably from a few realizations of the process. However, if we can find a basis in which the covariance operator is well approximated by a sparse matrix, it is possible to reduce substantially the variance by estimating only the (essentially) non-zero matrix elements. It is thus necessary to estimate from the data the basis in which the covariance operator is well approximated by a sparse matrix, as well as the non-zero matrix elements. A best basis search algorithm has been introduced to compress the covariance operator.

Locally stationary processes appear in many physical systems in which the mechanisms that produce random fluctuations change slowly in time or space. Over short time intervals, such processes can be approximated by a stationary one. This is the case for many components of speech signals. We have shown that the covariance operator of such processes is nearly diagonalized in an appropriate local cosine basis. A best basis algorithm was designed to search this best basis and estimate the compressed covariance matrix.

9 Technical Publications

Journal Publications

1. S. Mallat, S. Zhang, "Matching Pursuits With Time-Frequency Dictionaries", IEEE Transactions on Signal Processing, December 1993.

2. Wen Liang Hwang, S. Mallat "Characterization of Self-Similar Multifractals with Wavelet Maxima", *Applied and Computational Harmonic Analysis"*, vol. 1, p. 316-328, 1994.

3. G. Davis, S. Mallat, M. Avellaneda, "Adaptive Nonlinear Approximations", To appear in *Journal of Constructive Approximation.*

4. G. Davis, S. Mallat and Z. Zhang, "Adaptive Time-Frequency Decompositions", *SPIE Journal of Optical Engineering,* vol. 33, No. 7, p. 2183-2191, July 1994.

5. G. Davis, S. Mallat and M. Avelaneda, "Adaptive Greedy Approximations", to appear in Constructive Approximations.

6. S. Mallat,"Wavelets for a Vision", Proceeding of the IEEE, vol. 4, no. 4.

7. F.Bergeaud, S.Mallat, "Matching Pursuit: Adaptive representations of images and sounds", To appear in Computational and Applied Mathematics, Birkhaüser, Boston.

8. S. Mallat, G. Papanicolaou, Z. Zhang, "Adaptive Covariance Estimation of Locally Stationary Processes", invited paper in the Annals of Statistics.

9. S. Jaggi, W. Karl, S. Mallat, A. Willsky, "High Resolution Pursuit for Feature Extraction", submitted to Applied and Computational Harmonic Analysis.

10 Research Personnel and Ph.D Degree Awarded

Stephane Mallat, PI
Wen Liang Hwang, Ph.D degree in computer science, Courant Institute,August

1993, “Singularity Detection, Noise reduction and Multifractal Characterization using Wavelet”.
Zhifeng Zhang, Ph.D degree in mathematics, Courant Institute, August 1993, “Matching Pursuit”.
Geoff Davis, Ph.D degree in mathematics, Courant Institute, July 1994, “Chaos in Matching Pursuit”.
Francois Bergeaud, research scientist.